

# **Exposé de Mathieu Roche, *Cirad, TETIS***

## **Science des Données Textuelles**

L'analyse des masses de données, en particulier textuelles, nécessite l'utilisation de méthodes mêlant harmonieusement différentes disciplines comme l'informatique, les mathématiques ou les statistiques. L'ensemble de ces méthodes utiles pour traiter de telles données forme le socle de la "Science des Données Textuelles". Dans ce cadre, les approches de fouille de textes permettent de découvrir des connaissances utiles pour des experts issus généralement de différents domaines d'application (par exemple, veille épidémiologique, analyse de données de SHS, etc.). L'exposé présentera les fondements des approches et leur mise en oeuvre pour traiter les données textuelles hétérogènes à travers différents projets pluridisciplinaires.

---

## **Textual Data Science**

The analysis of large volumes of data, in particular textual data, requires the use of methods that combine different disciplines such as computer science, mathematics, or statistics. All these methods represent the basis of "Textual Data Science". In this context, text-mining approaches enable to discover knowledge useful for experts of different domains (e.g. Epidemiology, Humanities, etc.). The talk will present some approaches and their implementation in order to deal with heterogenous textual data in the context of multidisciplinary projects.



**LIRIMA Scientific Days 2017**  
**September 12th – 14th 2017**

**ENIT – Tunis**

**<http://lirima.inria.fr/events/2017-scientific-days/>**

The Scientific Days will be held from September 12 to September 14, 2017 at **ENIT** (École Nationale d'Ingénieurs de Tunis), located on the campus of El Manar Tunis University in Tunis.

See the program below.

**Practical information:**

- **Location:** salle Chaire UNESCO  
Campus of Farhat HACHED EL MANAR University in Tunis  
BP 37, LE BELVEDERE 1002 TUNIS
- **Accommodation:** rooms are booked at Hotel Ariha for the participants who asked for accommodation in their registration form.  
<http://www.hotel-ariha.com/index.php>
- **Access:** the participants will be welcomed at the airport by our colleagues from ENIT.  
A free shuttle will pick up the participants from the hotel to the Campus every day.
- **Contacts:**  
Amel Ben Abda – ENIT : mobile : 00 216 24 383 661  
Nabil Gmati – ENIT : fixe : 00 216 71 871 022 – mobile : 00 216 20 233 861  
Laura Norcy – Inria : mobile : + 33 (0)6 86 38 93 62

<b>Tuesday September 12<sup>th</sup> - salle Chaire UNESCO</b>	
9:00-9:30	<b>Welcome and introduction</b>
9:30-12:30	<b>Presentation of the Associate Teams</b>
9:30-10:15	SIMERGE
10:15-10:45	EPITAG
10:45-11:15	<i>Coffee break</i>
11:15-12:00	IoT4D
12:00-12:30	AGRINET
12:30-13:30	<b>Lunch</b>
13:30-16:30	<b>Presentation of the Associate Teams</b>
13:30-14:15	MOHA
14:15-15:00	EPICARD

15:00-15:15	<b>Short break</b>
15:15-16:00	MOREWAIS
16:00-16:30	FAST
16:30-17:00	<b>Coffee break</b>
17:00-18:30	<b>Meeting with the Scientific Council and Associate Teams members</b>
20:00	<b>Joint Dinner</b>
<b>Wednesday September 13<sup>th</sup> - salle de conférence Mokhtar Laatiri</b>	
9:00-17:00	<b>Workshop "Data Sciences"</b>
9:00-10:00	<i>Mohamed Masmoudi, Univ. Paul Sabatier, Toulouse - " Deep learning: Nothing is left to chance "</i>
10:00-10:15	<b>Coffee break</b>
10:15-11:15	<i>Mathieu Roche, Cirad – " <u>Textual Data Science</u> "</i>
11:15-11:30	<b>Coffee break</b>
11:30-12:30	<i>Jean-Daniel Zücker, IRD – " <u>Machine Learning Challenges in Metagenomics Data Analysis</u> "</i>
12:30-13:30	<b>Lunch</b>
13:30-14:00	<i>Riaan Wolhuter, Stellenbosch University &amp; Nathalie Mitton, Inria – " <u>A digital VHF radio communications system for Zinave National Park, Mozambique: a case study</u> "</i>
14:00-14:30	<i>Olivier Corby, Inria – " <u>STTL, un langage de transformation pour les graphes RDF du Web de données liées</u> "</i>
14:30-15:00	<i>Nejib Zemzemi, Inria – " Examples of challaging questions in computational cardiology: High dimension, multiphysics and heterogeneous data "</i>
15:00-15:15	<b>Coffee break</b>
15:15-16:15	<i>Lilia Sfaxi, INSAT – " <u>Thinking Big</u> "</i>
16:15-16:30	<b>Coffee break</b>
16:30-17:30	<i>Claude Kirchner, Inria – " <u>OpenData and openSciences: Opportunities and Responsible Research and Innovation</u> "</i>
17:30-17:45	<b>Coffee break</b>
17:45-18:45	<b>Panel with managers of Tunisian companies:</b> <ul style="list-style-type: none"> <li>• <i>Karim Beguir, Instadeep – Using supercomputers in AI</i></li> <li>• <i>Samy Achour, Integration Objects – Data Sciences dans l'industrie 4.0</i></li> <li>• <i>Hédi Zaher &amp; Houda Bakir, Datavora – Big Data et Data Science: game changers du E-Commerce</i></li> </ul>
18:45-19:30	<b>Cocktail</b>
20:30	<b>Dinner</b>
<b>Thursday September 14<sup>th</sup> - salle Chaire UNESCO</b>	
9:30-12:00	<b>Meeting between the Steering Committee and the Scientific Council</b>
12:00-13:30	<b>Lunch</b>
13:30	Departure